

Article

Optimization of Aerosol Model Selection for TROPOMI/S5P

Lanlan Rao ^{1,2} , Jian Xu ^{1,*} , Dmitry S. Efremenko ¹ , Diego G. Loyola ¹  and Adrian Doicu ¹

¹ Remote Sensing Technology Institute, German Aerospace Center (DLR), 82234 Oberpfaffenhofen, Germany; lanlan.rao@dlr.de (L.R.); dmitry.efremenko@dlr.de (D.S.E.); diego.loyola@dlr.de (D.G.L.); adrian.doicu@dlr.de (A.D.)

² Department of Aerospace and Geodesy, Technische Universität München, 80331 Munich, Germany

* Correspondence: jian.xu@dlr.de; Tel.: +49-8153-28-3353

† Current address: National Space Science Center, Chinese Academy of Sciences, Beijing 100190, China.

Abstract: To retrieve aerosol properties from satellite measurements, micro-physical aerosol models have to be assumed. Due to the spatial and temporal inhomogeneity of aerosols, choosing an appropriate aerosol model is an important task. In this paper, we use a Bayesian algorithm that takes into account model uncertainties to retrieve the aerosol optical depth and layer height from synthetic and real TROPOMI O₂A band measurements. The results show that in case of insufficient information for an appropriate micro-physical model selection, the Bayesian algorithm improves the accuracy of the solution.

Keywords: model selection; aerosol retrievals; TROPOMI/S5P



Citation: Rao, L.; Xu, J.; Efremenko, D.S.; Loyola, D.G.; Doicu, A. Aerosol Model Selection for TROPOMI/S5P. *Remote Sens.* **2021**, *13*, 2489. <https://doi.org/10.3390/rs13132489>

Academic Editor: Luke Knibbs

Received: 14 May 2021

Accepted: 22 June 2021

Published: 25 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Aerosols affect the Earth's climate directly by disturbing the Earth's radiation budget and indirectly by altering cloud processes. To better understand the role of aerosols in the Earth's climate, it is important to observe concentrations and properties of aerosols. Satellite sensors provide long-term measurements that can effectively monitor aerosol information on both regional and global scales.

The information on the aerosol optical depth can be retrieved from the data provided by satellite sensors, such as the Advanced Very High Resolution Radiometer (AVHRR) [1], the Moderate Resolution Imaging Spectroradiometer (MODIS) [2], the Visible Infrared Imaging Radiometer (VIIRS) [3], and the Advanced Himawari Imager (AHI) [4], helping to understand the temporal and spatial distribution characteristics of atmospheric aerosols. Aerosol height information can be retrieved from (i) multi-angle instruments, e.g., the Multi-angle Imaging Spectroradiometer (MISR) [5], and the Advanced Along-Track Scanning Radiometer (AATSR) [6]; (ii) polarization measurements, e.g., the Polarization and Directionality of the Earth's Reflectances (POLDER) [7]; and (iii) measurements in the oxygen absorption band, e.g., the TROPospheric Monitoring Instrument (TROPOMI) [8]. A combination of multi-angle and polarization observations [9] can also provide information of micro-physical parameters such as particle size distribution and refractive index.

However, the information that can be retrieved from space is quite limited. To retrieve the aerosol parameters, aerosol models characterizing the micro-physical properties have to be assumed. Aerosol properties exhibit high spatial inhomogeneity because of various origins and complex processes during transportation in the atmosphere. Aerosol particles are originated from both natural processes (such as wind-blown desert dust and sea salt, wild forest fire, and volcano eruption) and anthropogenic activities (such as industrial activities, artificial vegetation fire, and fossil fuel combustion). The selection of a suitable aerosol model in the retrieval algorithm relies on the knowledge of emission sources.

There are several databases and sets of aerosol models portraying the aerosol micro-physical properties on a global scale. The Optical Properties of Aerosols and Clouds (OPAC) database [10] describes the size distribution and spectral refractive index of 10 aerosol

components under different humidities. These components can form various aerosol types through internal mixture. The dark-target algorithm of the Moderate-Resolution Imaging Spectroradiometer (MODIS) characterizes a set of aerosol models and provides global distributions of aerosol types for different seasons based on a cluster analysis of the AERONET climatology [2,11]. The OMI near-UV (OMAERUV) algorithm and the multi-wavelength algorithm (OMAERO) consider several major aerosol types which are split into different aerosol models. The selection of an aerosol model is based on spectral and geographic considerations [12]. A chemical transport model, such as the Goddard Chemistry Aerosol Radiation and Transport (GOCART) model also supplies distributions of different aerosol types [13,14]. Besides, a number of studies coping with classification of aerosol types based on satellite remote sensing were carried out, see, e.g., in [15–19].

In standard retrieval algorithms, an aerosol model is chosen from a set of candidate models, and the retrieval is performed as if the selected model reflects the real scenario. In general, model selection is not a trivial task because for a given measurement, several models may fit the data equally well. The Bayesian approach and, in particular, the Bayesian model selection and model averaging Hoeting et al. [20], is a statistical method using measurement data to select the best fitting models from a set of candidate models without any prior seasonal or geographical information. The Bayesian method provides a posteriori probability densities for given models, also known as model evidences. In the Bayesian model selection, we select the model with the highest evidence, while in the Bayesian model averaging, we combine the retrieval results corresponding to different candidate models weighted by their evidences. Määttä et al. [21] introduced the Bayesian approach into the aerosol model selection of the OMAERO algorithm, Kauppi et al. [22] applied the Bayesian approach to real data of OMI, while Sasi et al. [23,24] applied the Bayesian approach to EPIC (Earth Polychromatic Imaging Camera) [25] synthetic measurements.

In this paper, for the first time, we use the Bayesian approach to jointly retrieve the aerosol optical depth and aerosol layer height from TROPOMI/S5P (Sentinel-5 Precursor) [26] measurements in the O₂A band (758–771 nm). TROPOMI is a hyperspectral instrument on board the Copernicus Sentinel-5 Precursor satellite launched on 13 October 2017, measuring the solar radiance backscattered by atmosphere and Earth's surface in the ultraviolet (UV), visible (VIS), near-infrared (NIR), and shortwave infrared (SWIR) spectral ranges. The aerosol parameters are retrieved using NIR measurements with a spectral resolution of ~0.45 nm. As the first atmospheric monitoring mission within the Copernicus program, TROPOMI has a very high spatial resolution of $3.5 \times 7 \text{ km}^2$ ($3.5 \times 5.5 \text{ km}^2$ since 6 August 2019), as compared with its predecessors. In particular, the spectra in O₂A band (758–771 nm) provides a way to retrieve the aerosol height information. The physical principle of aerosol height detection in O₂A band lies on the fact that the aerosol layer attenuate the reflection of solar radiance by the lower atmosphere at high oxygen absorption wavelengths. This attenuation decreases as the decline of oxygen absorption coefficient. To our best knowledge, currently no satellite passive sensor except TROPOMI can provide official product of aerosol height information.

The paper is organized as follows. In Section 2, we review the Bayesian model selection approach and discuss its application to aerosol retrievals. Section 3 describes the sets of aerosol models used in our numerical analysis. The accuracy of the Bayesian model selection approach is analyzed in Section 4 for synthetic measurements and in Section 5 for real data over a wild fire scene in South Africa.

2. Methodology

We have developed a retrieval algorithm dedicated to satellite remote sensing of aerosol and cloud parameters. The physics-based retrieval algorithm comprises a forward model calculating radiative transfer of electromagnetic radiation through a planetary atmosphere and an inversion module solving a nonlinear minimization problem. In the forward model, the radiative transfer calculation depends on the discrete ordinate method with

matrix exponential. To speed up the computation in the oxygen absorption band from sensors (e.g., TROPOMI) with very high spectral resolution, we have implemented acceleration techniques like the telescoping technique [27,28], the false discrete ordinate approach, the correlated k -distribution method [29], and the principal components analysis [30,31]. The inversion is performed by the means of Tikhonov regularization with optimal strategies for constructing the regularization parameter and matrix [32,33]. For further details about the retrieval algorithm and its forward model, we refer to the works in [34–36].

In this study, the aerosol optical depth τ and the layer height h are retrieved in the O₂A band (758–771 nm). The retrieval algorithm can deal with four types of aerosol profiles: Gaussian, exponential decay, elevated box, and a combination of exponential decay and ground box. To simplify the analysis, the aerosol layer is assumed to be homogeneous, spreading evenly from near surface to the top aerosol layer height h . Considering N_m aerosol models, the retrieval of the state vector $\mathbf{x} = [\tau, h]$ is an inverse problem relying on the solution of the nonlinear equation

$$\mathbf{y}^\delta = \mathbf{F}_m(\mathbf{x}) + \delta_m, \quad (1)$$

where \mathbf{y}^δ is the measurement vector, $\mathbf{F}_m(\mathbf{x})$ is the forward model corresponding to the aerosol model m with $m = 1, \dots, N_m$, $\delta_m = \delta_{\text{mes}} + \delta_{\text{aerm}}$ the total data error vector, δ_{mes} the measurement error vector, and δ_{aerm} the aerosol model error vector, i.e., the error due to an inadequate aerosol model. In our analysis, $\mathbf{F}_m(\mathbf{x})$ is the vector of the log of the simulated radiances corresponding to aerosol model m , i.e., $[\mathbf{F}_m(\mathbf{x})]_i = \ln I_m(\lambda_i, \mathbf{x})$, where $\{\lambda_i\}_{i=1}^M$ is a set of M wavelengths in the considered spectral domain.

The data model (1) is transformed into a model with white noise by using the prewhitening technique. The procedure is as follows. Assuming that

1. δ_{mes} is a Gaussian random vector with zero mean and covariance matrix $\mathbf{C}_{\text{mes}} = \sigma_{\text{mes}}^2 \mathbf{C}_{\text{mes}}$, where σ_{mes}^2 is the measurement error variance and \mathbf{C}_{mes} a normalized measurement error covariance matrix;
2. δ_{aerm} is a Gaussian random vector with zero mean and covariance matrix $\mathbf{C}_{\text{aerm}} = \sigma_{\text{aerm}}^2 \mathbf{I}_M$, where σ_{aerm}^2 is the aerosol model error variance and \mathbf{I}_M the identity matrix; and
3. δ_{mes} and δ_{aerm} are independent random vectors,

we deduce that δ_m is also a Gaussian random vector with zero mean and covariance matrix $\mathbf{C}_{\delta m} = \mathbf{C}_{\text{mes}} + \mathbf{C}_{\text{aerm}} = \sigma_m^2 \mathbf{C}_{\delta m}$, where $\sigma_m^2 = \sigma_{\text{mes}}^2 + \sigma_{\text{aerm}}^2$ is the data error variance and $\mathbf{C}_{\delta m} = w_{\text{mes}m} \mathbf{C}_{\text{mes}} + (1 - w_{\text{mes}m}) \mathbf{I}_M$ with $w_{\text{mes}m} = \sigma_{\text{mes}}^2 / \sigma_m^2$, a normalized data error covariance matrix. In this context, the scaled data model reads as

$$\bar{\mathbf{y}}^\delta = \bar{\mathbf{F}}_m(\mathbf{x}) + \bar{\delta}_m, \quad (2)$$

where $\bar{\mathbf{y}}^\delta = \mathbf{P} \mathbf{y}^\delta$, $\bar{\mathbf{F}}_m(\mathbf{x}) = \mathbf{P} \mathbf{F}_m(\mathbf{x})$, $\bar{\delta}_m = \mathbf{P} \delta_m$, and $\mathbf{P} = \mathbf{C}_{\delta m}^{-1/2}$ is a scaling matrix. As $\bar{\mathbf{C}}_{\delta m} = \mathcal{E}\{\bar{\delta}_m \bar{\delta}_m^T\} = \sigma_m^2 \mathbf{I}_M$, it is readily seen that $\bar{\delta}_m \sim \mathcal{N}(\mathbf{0}, \bar{\mathbf{C}}_{\delta m} = \sigma_m^2 \mathbf{I}_M)$, where the notation $\mathcal{N}(\mathbf{x}_{\text{mean}}, \mathbf{C}_x)$ stands for a normal distribution with mean \mathbf{x}_{mean} and covariance matrix \mathbf{C}_x . In a stochastic setting, we assume that $\mathbf{x} \sim \mathcal{N}(\mathbf{x}_a, \mathbf{C}_x)$, where \mathbf{x}_a is the a priori state vector, the best beforehand estimate of the solution, $\mathbf{C}_x = \sigma_x^2 \mathbf{C}_x$ is the a priori covariance matrix, and σ_x^2 the a priori state variance. Defining the regularization matrix \mathbf{L} and the regularization parameter α through the relations $\mathbf{C}_x^{-1} = \mathbf{L}^T \mathbf{L}$ and $\alpha = \sigma_m^2 / \sigma_x^2$, respectively, we express the a priori covariance matrix as $\mathbf{C}_x = \sigma_m^2 (\alpha \mathbf{L}^T \mathbf{L})^{-1}$.

The scaled nonlinear Equation (2) is solved by means of a Bayesian approach. The key quantity in this approach is the a posteriori density $p(\mathbf{x} | \bar{\mathbf{y}}^\delta, m)$, which represents the conditional probability density of \mathbf{x} given the data $\bar{\mathbf{y}}^\delta$ and the aerosol model m . According to Bayes' theorem, the a posteriori density is given by

$$p(\mathbf{x} | \bar{\mathbf{y}}^\delta, m) = \frac{p(\bar{\mathbf{y}}^\delta | \mathbf{x}, m) p(\mathbf{x} | m)}{p(\bar{\mathbf{y}}^\delta | m)},$$

where $p(\mathbf{x} | m)$ is the a priori density, i.e., the conditional probability density of \mathbf{x} given the aerosol model m before performing the measurement $\bar{\mathbf{y}}^\delta$, $p(\bar{\mathbf{y}}^\delta | \mathbf{x}, m)$ the likelihood density, i.e., the conditional probability density of $\bar{\mathbf{y}}^\delta$ given the state \mathbf{x} and the aerosol model m , and

$$p(\bar{\mathbf{y}}^\delta | m) = \int p(\mathbf{x}, \bar{\mathbf{y}}^\delta | m) d\mathbf{x} = \int p(\bar{\mathbf{y}}^\delta | \mathbf{x}, m) p(\mathbf{x} | m) d\mathbf{x}, \quad (3)$$

the marginal likelihood density. Although, in the Bayesian parameter estimation, the marginal likelihood density $p(\bar{\mathbf{y}}^\delta | m)$ plays the role of a normalization constant and is usually ignored, this probability density is of particular importance in the Bayesian model selection. For $\mathbf{x} \sim N(\mathbf{x}_a, \mathbf{C}_x = \sigma_m^2(\alpha \mathbf{L}^T \mathbf{L})^{-1})$ and $\bar{\delta}_m \sim N(\mathbf{0}, \bar{\mathbf{C}}_{\delta m} = \sigma_m^2 \mathbf{I}_M)$, the Bayes' formula gives

$$p(\mathbf{x} | \bar{\mathbf{y}}^\delta, m) \propto \exp\left[-\frac{1}{2} V_\alpha(\mathbf{x} | \bar{\mathbf{y}}^\delta, m)\right], \quad (4)$$

where

$$V_\alpha(\mathbf{x} | \bar{\mathbf{y}}^\delta, m) = \frac{1}{\sigma_m^2} \left[\|\bar{\mathbf{y}}^\delta - \bar{\mathbf{F}}_m(\mathbf{x})\|^2 + \alpha \|\mathbf{L}(\mathbf{x} - \mathbf{x}_a)\|^2 \right] \quad (5)$$

is the a posteriori potential. Consequently, the maximum a posteriori estimate $\hat{\mathbf{x}}_{m\alpha}^\delta$ is computed as

$$\hat{\mathbf{x}}_{m\alpha}^\delta = \arg \min_{\mathbf{x}} V_\alpha(\mathbf{x} | \bar{\mathbf{y}}^\delta, m). \quad (6)$$

In a deterministic setting, $\mathcal{F}_{m\alpha}(\mathbf{x}) = \sigma_m^2 V_\alpha(\mathbf{x} | \bar{\mathbf{y}}^\delta, m)$ is the Tikhonov function for the nonlinear equation $\bar{\mathbf{y}}^\delta = \bar{\mathbf{F}}_m(\mathbf{x})$ with the penalty term $\alpha \|\mathbf{L}(\mathbf{x} - \mathbf{x}_a)\|^2$ and the regularization parameter α . Thus, a regularized solution $\mathbf{x}_{m\alpha}^\delta$, which minimizes $\mathcal{F}_{m\alpha}(\mathbf{x})$, coincides with the maximum a posteriori estimate, i.e., $\mathbf{x}_{m\alpha}^\delta = \hat{\mathbf{x}}_{m\alpha}^\delta$. The computation of the regularized solution $\mathbf{x}_{m\alpha}^\delta$ in the framework of the method of Tikhonov regularization requires the knowledge of the optimal value of the regularization parameter $\hat{\alpha}$. Because in practice, this is a very challenging task, the nonlinear equation $\bar{\mathbf{y}}^\delta = \bar{\mathbf{F}}_m(\mathbf{x})$ is solved by means of the iteratively regularized Gauss–Newton method. This method provides an optimal value of the regularization parameter $\hat{\alpha}$ (i.e., the ratio of the data error variance σ_m^2 and the a priori state variance σ_x^2) and the corresponding regularized solution $\mathbf{x}_{m\hat{\alpha}}^\delta$.

For model comparison, the so-called relative evidence of the m th aerosol model $p(m | \bar{\mathbf{y}}^\delta)$ plays an important role. In view of Bayes' theorem, this conditional probability density is defined by

$$p(m | \bar{\mathbf{y}}^\delta) = \frac{p(\bar{\mathbf{y}}^\delta | m) p(m)}{p(\bar{\mathbf{y}}^\delta)} = \frac{p(\bar{\mathbf{y}}^\delta | m)}{\sum_{m=1}^{N_m} p(\bar{\mathbf{y}}^\delta | m)}, \quad (7)$$

where the last equality follows from the mean formula $p(\bar{\mathbf{y}}^\delta) = \sum_{m=1}^{N_m} p(\bar{\mathbf{y}}^\delta | m) p(m)$ and the assumption that all aerosol models are equally like, i.e., $p(m) = 1/N_m$. In terms of $p(m | \bar{\mathbf{y}}^\delta)$, the mean a posteriori density reads as

$$p_{\text{mean}}(\mathbf{x} | \bar{\mathbf{y}}^\delta) = \sum_{m=1}^{N_m} p(\mathbf{x} | \bar{\mathbf{y}}^\delta, m) p(m | \bar{\mathbf{y}}^\delta), \quad (8)$$

while the mean and the maximum solution estimates are defined by

$$\hat{\mathbf{x}}_{\text{mean}}^\delta = \sum_{m=1}^{N_m} \mathbf{x}_{m\hat{\alpha}}^\delta p(m | \bar{\mathbf{y}}^\delta), \quad (9)$$

and

$$\hat{\mathbf{x}}_{\text{max}}^\delta = \mathbf{x}_{m^* \hat{\alpha}}^\delta, \quad m^* = \arg \max_m p(m | \bar{\mathbf{y}}^\delta), \quad (10)$$

respectively. In Equation (9), the Bayesian model averaging technique is used to combine the individual solutions weighted by their evidences, while in Equation (8), the aerosol

model with the highest evidence is considered to be the best among all aerosol models involved. Intuitively, we expect that in practice, a linear combination of the retrieved parameters corresponding to different candidate models will better reproduce the real scenario than the retrieved parameters corresponding to an a priori selected model. In a stochastic setting, the relative evidence $p(m | \bar{\mathbf{y}}^\delta)$ can be computed via Equation (7) by using a linearization of the forward model around the solution and under the assumption that the data error variance σ_m^2 is known. In [23,24], estimates for σ_m^2 were obtained in the framework of the maximum marginal likelihood estimation [37–39] and the generalized cross-validation method [40,41]. In a deterministic setting, $p(m | \bar{\mathbf{y}}^\delta)$, regarded as a merit function characterizing the solution $\mathbf{x}_{m\hat{\alpha}}^\delta$, can be defined in terms of the marginal likelihood function or the generalized cross-validation function. In the latter case, the computational formula is

$$p(m | \mathbf{y}^\delta) = \frac{1/v(m)}{\sum_{m=1}^{N_m} 1/v(m)}. \quad (11)$$

where

$$v(m) = \frac{\|\mathbf{r}_{m\hat{\alpha}}^\delta\|^2}{[\text{trace}(\mathbf{I} - \hat{\mathbf{A}}_{m\hat{\alpha}})]^2} \quad (12)$$

is the generalized cross-validation function, $\mathbf{r}_{m\hat{\alpha}}^\delta = \bar{\mathbf{y}}^\delta - \bar{\mathbf{F}}_m(\mathbf{x}_{m\hat{\alpha}}^\delta)$ the nonlinear residual vector, $\hat{\mathbf{A}}_{m\hat{\alpha}} = \bar{\mathbf{K}}_{m\hat{\alpha}} \bar{\mathbf{K}}_{m\hat{\alpha}}^\dagger$ the influence matrix, $\bar{\mathbf{K}}_{m\hat{\alpha}}$ the Jacobian matrix, and $\bar{\mathbf{K}}_{m\hat{\alpha}}^\dagger$ the generalized inverse at the solution $\mathbf{x}_{m\hat{\alpha}}^\delta$.

3. Aerosol Models

Two sets of aerosol micro-physical models are used in our numerical analysis. The first set (Set 1) is taken from the MODIS dark-target algorithm [11] and includes the following four aerosol models:

1. non-absorbing (NONABS) aerosols, generated from fossil fuel combustion in urban-industrial areas;
2. moderately absorbing (MODABS) aerosols;
3. absorbing (ABS) aerosols, generated from biomass burning; and
4. desert dust (DUST), originated from desert and transported by wind.

The volume size distribution of each aerosol model is a bimodal log-normal distribution consisting of a fine and a coarse mode. The parameters of the size distribution (median radius, standard deviation, and volume of particles) and the complex refractive index, which depend on the aerosol optical depth, are illustrated in Table 1. The second set (Set 2) is taken from the OMAERO algorithm and includes the following five major aerosol types:

1. weakly absorbing (WA),
2. biomass burning (BB),
3. desert dust (DD),
4. marine (MA), and
5. volcanic (VO).

Each type is split into several aerosol models depending on their optical properties and particle size distribution. The parameters of the size distribution (median radius, standard deviation, and fraction of coarse mode) and the complex refractive index are shown in Table 2.

In the forward model, the scattering characteristics (e.g., the single scattering albedo, the phase function, and the asymmetry parameter) can be computed by the Mie theory in the case of spherical particles, and the null-field method with discrete sources in the case of spheroidal particles with a size parameter smaller than 50. For spheroidal particles with large size parameter, we use a precomputed database as given in [42]. In this study, the aerosol particles are assumed to be spherical for simplicity.

Table 1. Micro-physical properties of aerosols models of Set 1. Each model is composed of a fine and coarse mode. The median radius of the volume distribution r_v , standard deviation σ , the volume of particles V_0 , and complex refractive index m of each mode are listed in the table.

| Model | Mode | r_v (μm) | σ | $m = (\text{Re}(m), \text{Im}(m))$ | V_0 ($\mu\text{m}^3/\mu\text{m}^2$) |
|--------|--------|-------------------------|----------------------|---|---|
| NONABS | fine | $0.160 + 0.0434\tau$ | $0.364 + 0.1529\tau$ | $(1.42, 0.004 - 0.0015\tau)$ | $0.1718\tau^{0.821}$ |
| | coarse | $3.325 + 0.1411\tau$ | $0.759 + 0.0168\tau$ | | $0.0934\tau^{0.639}$ |
| MODABS | fine | $0.145 + 0.0203\tau$ | $0.374 + 0.1365\tau$ | $(1.43, 0.008 - 0.002\tau)$ | $0.1642\tau^{0.775}$ |
| | coarse | $3.101 + 0.3364\tau$ | $0.729 + 0.098\tau$ | | $0.1482\tau^{0.684}$ |
| ABS | fine | $0.134 + 0.0096\tau$ | $0.383 + 0.0794\tau$ | $(1.51, 0.02)$ | $0.1748\tau^{0.891}$ |
| | coarse | $3.448 + 0.9489\tau$ | $0.743 + 0.0409\tau$ | | $0.1043\tau^{0.682}$ |
| DUST | fine | $0.1416\tau^{-0.052}$ | $0.7561\tau^{0.148}$ | $(1.48\tau^{-0.021}, 0.0018\tau^{-0.08})$ | $0.087\tau^{1.026}$ |
| | coarse | 2.2 | $0.554\tau^{-0.052}$ | | $0.6786\tau^{1.057}$ |

Table 2. Micro-physical properties of aerosols models of Set 2. The median radius of the number size distribution r_{mod} , stand deviation σ , and complex refractive index m of two modes are listed in the table. w_{coarse} is the fraction of coarse mode in number concentration. The two lines of m for MA mod. abs. and MA abs. aerosol model are the refraction index of fine and coarse modes, respectively.

| Type | Model | Fine Mode | | Coarse Mode | | $m = (\text{Re}, \text{Im})$ | w_{coarse} |
|------|--------|------------------|------------|------------------|------------|------------------------------|-------------------------|
| | | r_{mod} | e^σ | r_{mod} | e^σ | | |
| WA | WA1101 | 0.078 | 1.499 | 0.497 | 2.160 | $(1.4, 5.0 \times 10^{-8})$ | (4.36×10^{-4}) |
| | WA1102 | 0.088 | 1.499 | 0.509 | 2.160 | $(1.4, 5.0 \times 10^{-8})$ | (4.04×10^{-4}) |
| | WA1103 | 0.137 | 1.499 | 0.567 | 2.160 | $(1.4, 5.0 \times 10^{-8})$ | (8.10×10^{-4}) |
| | WA1104 | 0.030 | 2.030 | 0.240 | 2.030 | $(1.4, 5.0 \times 10^{-8})$ | (1.53×10^{-2}) |
| | WA1201 | 0.078 | 1.499 | 0.497 | 2.160 | $(1.4, 4.0 \times 10^{-3})$ | (4.36×10^{-4}) |
| | WA1202 | 0.088 | 1.499 | 0.509 | 2.160 | $(1.4, 4.0 \times 10^{-3})$ | (4.04×10^{-4}) |
| | WA1203 | 0.137 | 1.499 | 0.567 | 2.160 | $(1.4, 4.0 \times 10^{-3})$ | (8.10×10^{-4}) |
| | WA1301 | 0.078 | 1.499 | 0.497 | 2.160 | $(1.4, 1.2 \times 10^{-2})$ | (4.36×10^{-4}) |
| | WA1302 | 0.088 | 1.499 | 0.509 | 2.160 | $(1.4, 1.2 \times 10^{-2})$ | (4.04×10^{-4}) |
| | WA1303 | 0.137 | 1.499 | 0.567 | 2.160 | $(1.4, 1.2 \times 10^{-2})$ | (8.10×10^{-4}) |
| BB | BB2101 | 0.074 | 1.537 | 0.511 | 2.203 | $(1.5, 1.0 \times 10^{-2})$ | (1.70×10^{-4}) |
| | BB2102 | 0.087 | 1.537 | 0.567 | 2.203 | $(1.5, 1.0 \times 10^{-2})$ | (2.06×10^{-4}) |
| | BB2103 | 0.124 | 1.537 | 0.719 | 2.203 | $(1.5, 1.0 \times 10^{-2})$ | (2.94×10^{-4}) |
| | BB2201 | 0.074 | 1.537 | 0.511 | 2.203 | $(1.5, 2.0 \times 10^{-2})$ | (1.70×10^{-4}) |
| | BB2202 | 0.087 | 1.537 | 0.509 | 2.203 | $(1.5, 2.0 \times 10^{-2})$ | (2.06×10^{-4}) |
| | BB2203 | 0.124 | 1.537 | 0.719 | 2.203 | $(1.5, 2.0 \times 10^{-2})$ | (2.94×10^{-4}) |
| | BB2102 | 0.087 | 1.537 | 0.509 | 2.203 | $(1.5, 3.0 \times 10^{-2})$ | (2.06×10^{-4}) |
| | BB2103 | 0.124 | 1.537 | 0.719 | 2.203 | $(1.5, 3.0 \times 10^{-2})$ | (2.94×10^{-4}) |
| DD | BB2101 | 0.074 | 1.537 | 0.511 | 2.203 | $(1.5, 3.0 \times 10^{-2})$ | (1.70×10^{-4}) |
| | DD3101 | 0.042 | 1.697 | 0.670 | 1.806 | $(1.53, 4.0 \times 10^{-3})$ | (4.35×10^{-3}) |
| | DD3102 | 0.052 | 1.697 | 0.670 | 1.806 | $(1.53, 4.0 \times 10^{-3})$ | (4.35×10^{-3}) |
| | DD3201 | 0.042 | 1.697 | 0.670 | 1.806 | $(1.53, 1.0 \times 10^{-2})$ | (4.35×10^{-3}) |
| | DD3202 | 0.052 | 1.697 | 0.670 | 1.806 | $(1.53, 1.0 \times 10^{-2})$ | (4.35×10^{-3}) |

Table 2. Cont.

| Type | Model | Fine Mode | | Coarse Mode | | $m = (\text{Re}, \text{Im})$ | w_{coarse} |
|------|--------------|------------------|--------------|------------------|--------------|--|-------------------------|
| | | r_{mod} | e^{σ} | r_{mod} | e^{σ} | | |
| MA | MA mod. abs. | 0.030 | 2.030 | 0.240 | 2.030 | $(1.4, 4.0 \times 10^{-3})$ $(1.4, 5.0 \times 10^{-8})$ | (1.55×10^{-4}) |
| | MA abs. | 0.030 | 2.030 | 0.240 | 2.030 | $(1.4, 1.2 \times 10^{-2})$ $(1.4, 5.0 \times 10^{-8})$ | (1.55×10^{-4}) |
| VO | VO4101 | 0.230 | 0.800 | 0.240 | 2.030 | $(1.45, 7.5 \times 10^{-8})$ | 0.5 |

Both sets of aerosol models have been widely used in satellite remote sensing of aerosol properties and are representative for characterizing aerosol microphysical properties. According to the EPIC experiment in [24], Set 2 slightly outperformed in the retrieval outcome and can be suggested as a proper choice.

4. Tests with Synthetic Data

In this section, we analyze the accuracy of the Bayesian model selection algorithm for synthetic measurements.

4.1. Test 1

In the first test example, synthetic measurement spectra are simulated for each aerosol model included in Set 1 ($m_e = \text{NONABS}, \text{MODABS}, \text{ABS}, \text{DUST}$), and for each measurement, all aerosol models from Set 1 are considered in the retrieval. Thus, the retrieval algorithm has the possibility of identifying the correct aerosol model. The exact aerosol optical depths and layer heights to be retrieved are

$$\tau_e = 0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75, 2 \quad (13)$$

and

$$h_e = 1.5, 2.5, 3.5, 4.5, 5.5, 6.5, 7.5, 8.5, 9.5 \text{ km}, \quad (14)$$

respectively. The a priori values, which coincide with the initial guesses, are $\tau_a = 2.0$ and $h_a = 2 \text{ km}$, and a Lambertian surface with an albedo of 0.05 being assumed. The solar zenith, viewing zenith, and relative azimuth angles are $\theta_o = 60^\circ$, $\theta_v = 0^\circ$, and $\Delta\phi = 180^\circ$. For the exact solution $\mathbf{x}_e = [\tau_e, h_e]$, the accuracy of the solution estimates is quantified through the relative errors

$$\varepsilon_{\text{mean}}^{\tau}(\tau_e, h_e) = \frac{|\tau_{\text{mean}} - \tau_e|}{\tau_e} \text{ and } \varepsilon_{\text{mean}}^h(\tau_e, h_e) = \frac{|h_{\text{mean}} - h_e|}{h_e} \quad (15)$$

corresponding to (cf. Equation (9)) $\hat{\mathbf{x}}_{\text{mean}}^{\delta} = [\tau_{\text{mean}}, h_{\text{mean}}]$ and

$$\varepsilon_{\text{max}}^{\tau}(\tau_e, h_e) = \frac{|\tau_{\text{max}} - \tau_e|}{\tau_e} \text{ and } \varepsilon_{\text{max}}^h(\tau_e, h_e) = \frac{|h_{\text{max}} - h_e|}{h_e} \quad (16)$$

corresponding to (cf. Equation (10)) $\hat{\mathbf{x}}_{\text{max}}^{\delta} = [\tau_{\text{max}}, h_{\text{max}}]$.

In Figures 1 and 2, we illustrate the variations of the relative errors with respect to the exact aerosol layer height h_e for $\tau_e = 0.5$ (i.e., $\varepsilon_{\text{mean,max}}^{\tau,h}(\tau_e = 0.5, h_e)$), and the aerosol optical depth τ_e for $h_e = 3.5 \text{ km}$ (i.e., $\varepsilon_{\text{mean,max}}^{\tau,h}(\tau_e, h_e = 3.5 \text{ km})$), respectively. The following conclusions can be drawn:

1. The relative errors corresponding to the maximum solution estimate ($\varepsilon_{\text{max}}^{\tau}$ and $\varepsilon_{\text{max}}^h$) are considered to be acceptable according to the scientific requirements defined in the pre-launch characterization tests and significantly smaller than those corresponding to the mean solution estimate ($\varepsilon_{\text{mean}}^{\tau}$ and $\varepsilon_{\text{mean}}^h$). Thus, the retrieval algorithm can recognize correctly the exact aerosol model.

2. For the maximum solution estimate, the retrieved aerosol optical depth achieves a higher accuracy than the retrieved aerosol layer height.
3. Different aerosol models could have similar a posteriori densities as the inversion process is not ideally perfect. An inappropriate aerosol model may occasionally be identified, which can result in unexpected errors ($\tau_e = 1, 1.25$).

The relative errors $\varepsilon_{\max}^{\tau}(\tau_e, h_e = 3.5 \text{ km})$ and $\varepsilon_{\max}^h(\tau_e, h_e = 3.5 \text{ km})$ attain their highest values (of about 0.22 and 0.016, respectively) for $m_e = \text{MODABS}$ and $\tau_e = 1, 1.25$. To explain this result, in Figure 3 we plot the a posteriori densities $p(\mathbf{x} = [\tau, h] | \bar{\mathbf{y}}^{\delta}, m)$ for $m = \text{NONABS}, \text{MODABS}, \text{ABS}, \text{DUST}$, and the mean a posteriori densities $p_{\text{mean}}(\mathbf{x} = [\tau, h] | \bar{\mathbf{y}}^{\delta})$ in the case $m_e = \text{MODABS}$, $\tau_e = 1.25$ and $h_e = 3.5 \text{ km}$. The plots indicate that a posterior density for $m = \text{DUST}$ is of similar height and width to that for $m = \text{MODABS}$; the maximum solution is achieved at $m = \text{MODABS}$, i.e., the DUST appears to be the model with the highest evidence.

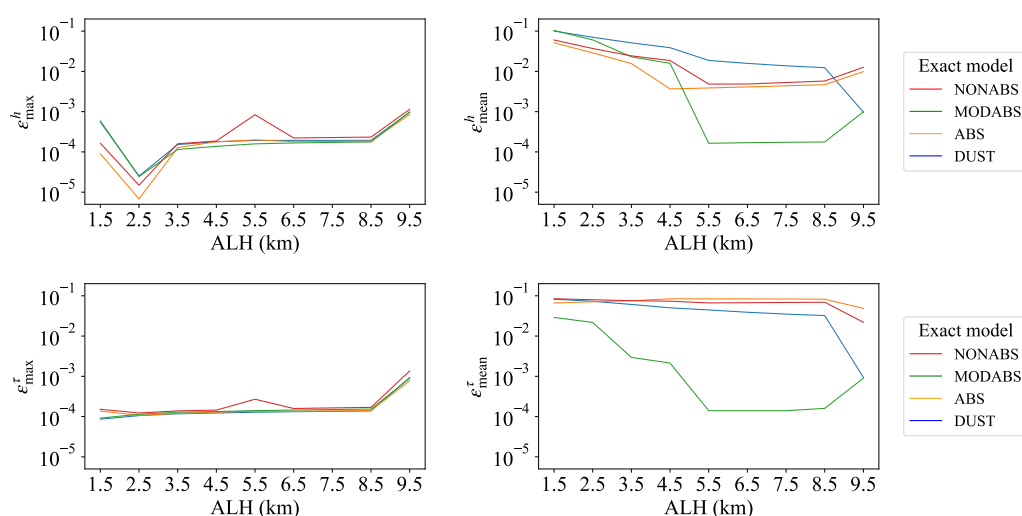


Figure 1. Relative errors $\varepsilon_{\text{mean}, \max}^{\tau, h}(\tau_e = 0.5, h_e)$ for the aerosol models from Set 1. ALH represents the aerosol layer height.

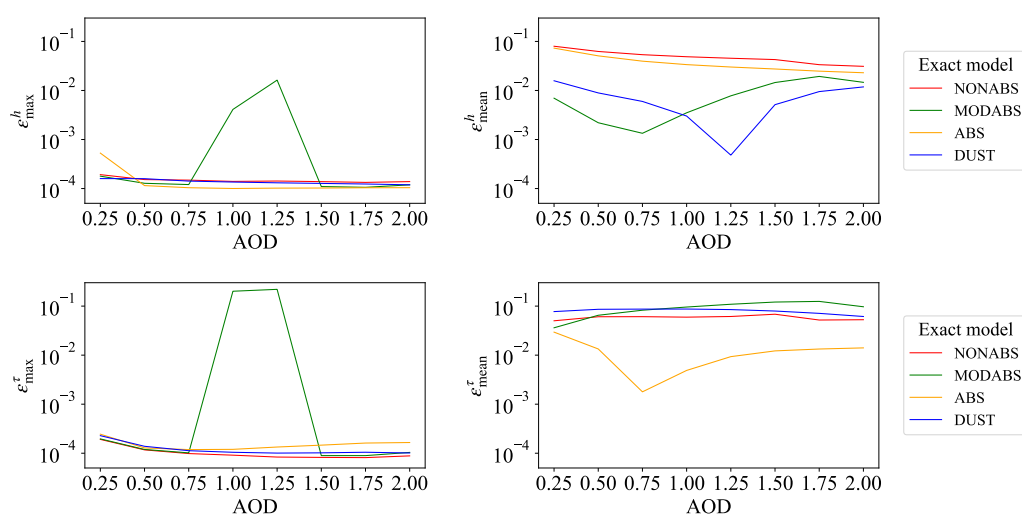


Figure 2. Relative errors $\varepsilon_{\text{mean}, \max}^{\tau, h}(\tau_e, h_e = 3.5 \text{ km})$ for the aerosol models from Set 1. AOD represents the aerosol optical depth.

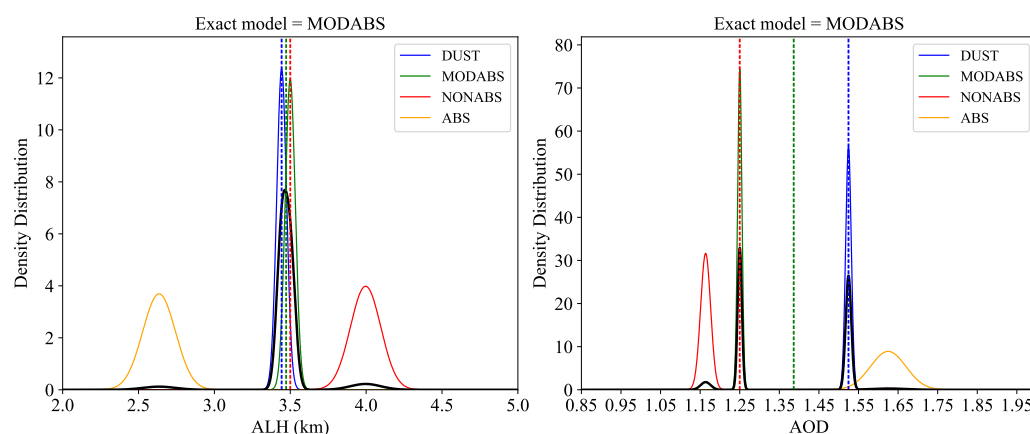


Figure 3. The a posteriori densities $p(\mathbf{x} = [\tau, h] | \bar{\mathbf{y}}^\delta, m)$ for $m = \text{NONABS}, \text{MODABS}, \text{ABS},$ and DUST , and the mean a posteriori densities $p_{\text{mean}}(\mathbf{x} = [\tau, h] | \bar{\mathbf{y}}^\delta)$ in the case $m_e = \text{MODABS}$, $\tau_e = 1.25$, and $h_e = 3.5$ km. The black curve indicates the mean a posterior density. In each plot, the red vertical dashed line corresponds to the exact values to be retrieved (τ_e, h_e), the blue vertical dashed line to the maximum solution estimate ($\tau_{\text{max}}, h_{\text{max}}$), and the green dashed line to the mean solution estimates ($\tau_{\text{mean}}, h_{\text{mean}}$).

4.2. Test 2

In the second test example, synthetic measurement spectra are produced for each aerosol model included in Set 1 ($m_e = \text{NONABS}, \text{MODABS}, \text{ABS}, \text{DUST}$), and for each measurement all aerosol models from Set 2 are considered in the retrieval. The mean solution estimate and the mean a posteriori density are computed for the first 10 aerosol models with the highest evidence.

The variations of the relative errors with respect to the exact aerosol layer height h_e for $\tau_e = 0.5$ and the aerosol optical depth τ_e for $h_e = 3.5$ km are illustrated in Figures 4 and 5, respectively. The plots indicate that

1. the relative errors are larger than those in the first test example,
2. the relative errors corresponding to the maximum solution estimate ($\epsilon_{\text{max}}^\tau$ and ϵ_{max}^h) and the mean solution estimate ($\epsilon_{\text{mean}}^\tau$ and ϵ_{mean}^h) are comparable, and
3. on average, the retrieved aerosol layer height obtains a higher accuracy than the retrieved aerosol optical depth.

The mean a posteriori densities $p_{\text{mean}}(\mathbf{x} = [\tau, h] | \bar{\mathbf{y}}^\delta)$ are shown in Figures 6 and 7 for $\tau_e = 0.5$, $h_e = 3.5$ km, and all exact aerosol models $m_e = \text{NONABS}, \text{MODABS}, \text{ABS},$ and DUST . The following conclusions could be made:

1. h_{mean} and h_{max} are both not too far from h_e ; thus, for aerosol layer height retrieval, the maximum solution estimate and the mean solution estimate ($\epsilon_{\text{mean}}^\tau$ and ϵ_{mean}^h) have similar accuracies;
2. τ_{mean} is relatively closer to τ_e than τ_{max} ; thus, for aerosol optical depth retrieval, the mean solution estimate performs better than the maximum solution estimate;
3. aerosol layer height retrievals have wide a posteriori densities that cover the exact layer height; and
4. aerosol optical depth retrievals have multi-peaked densities, in which the exact optical depth does not have the highest probability.

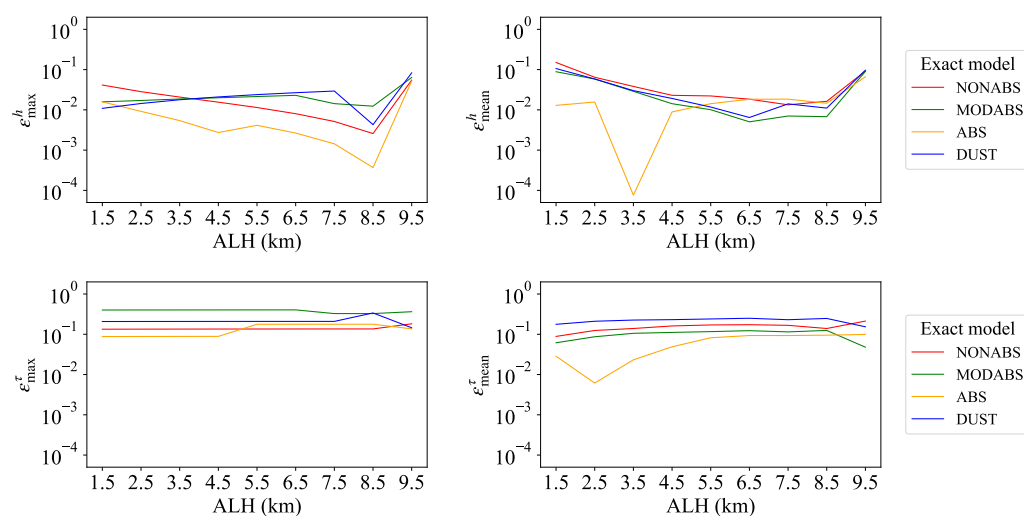


Figure 4. Relative errors $\varepsilon_{\text{mean,max}}^{\tau,h}(\tau_e = 0.5, h_e)$ for the aerosol models from Set 2.

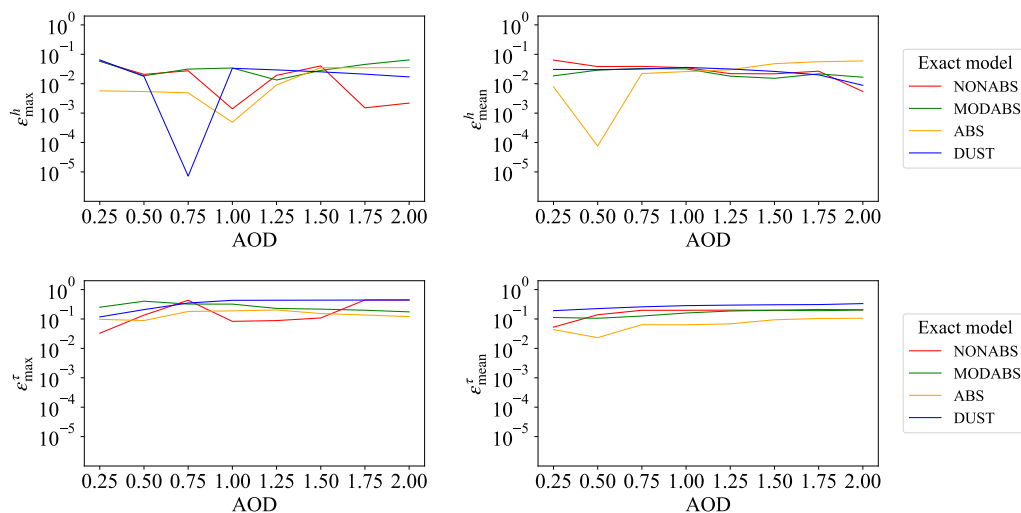


Figure 5. Relative errors $\varepsilon_{\text{mean,max}}^{\tau,h}(\tau_e, h_e = 3.5 \text{ km})$ for the aerosol models from Set 2.

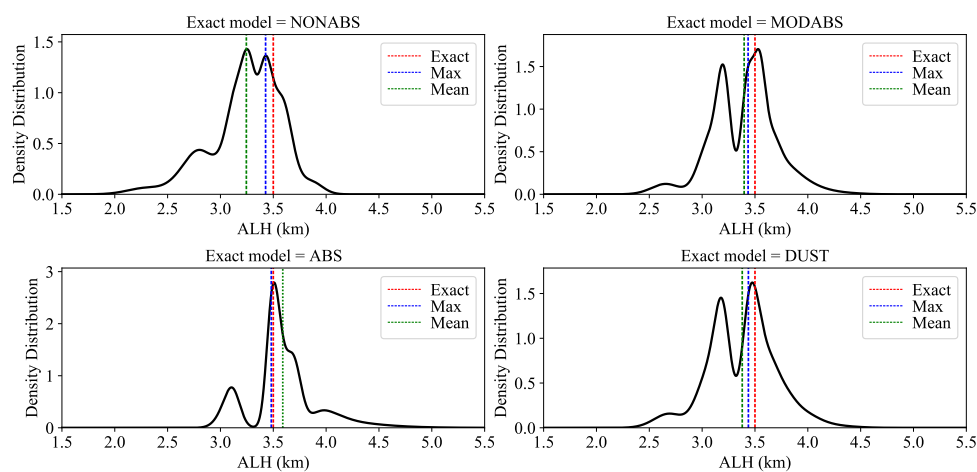


Figure 6. The mean a posteriori densities $p_{\text{mean}}(h | \bar{y}^{\delta})$ for $m_e = \text{NONABS, MODABS, ABS, DUST}$, $\tau_e = 0.5$, and $h_e = 3.5 \text{ km}$. In each plot, the red vertical dashed line correspond to the exact values to be retrieved (τ_e, h_e) , the blue vertical dashed line to the maximum solution estimate $(\tau_{\text{max}}, h_{\text{max}})$, and the green dashed line to the mean solution estimates $(\tau_{\text{mean}}, h_{\text{mean}})$.

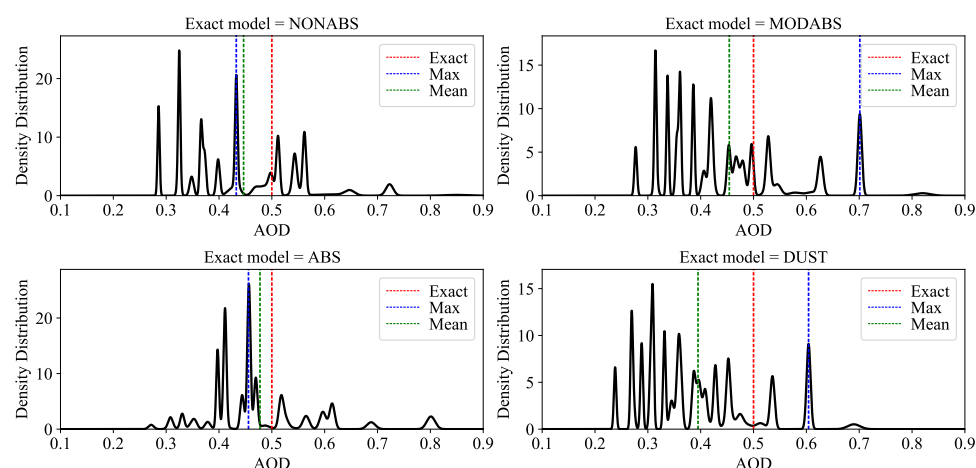


Figure 7. The same as in Figure 6 but for the mean a posteriori densities $p_{\text{mean}}(\tau | \bar{y}^{\delta})$.

5. Case Study with TROPOMI Data

To test the performance of the retrieval algorithm on real TROPOMI data, we chose a wild fire scene in South Africa and considered the measurements recorded on 4–5 July 2019. As can be seen from the respective VIIRS images, the wild fire smoke on 4 July 2019 (Figure 8a) traveled beyond the coastline and extended over the ocean, so that the smoke on 5 July 2019 (Figure 8b) was thinner. Regional studies of aerosol optical/microphysical properties during biomass burning can be found in [43–45].

The aerosol models included in Sets 1 and 2 are used in the retrieval. To decrease the retrieval uncertainty caused by unrealistic surface properties, the geometry-dependent effective Lambertian equivalent reflectivity (GE_LER) products [46] are used. The ground pixels with cloud fraction larger than 0.15 are excluded for this analysis, in which case, the scene is assumed to be cloud free so that we can retrieve valid aerosol information on sufficient number of pixels without significant impact by cloud contamination.

Figure 9 shows the aerosol model with the highest evidence from Set 1 as well as the aerosol type containing the aerosol model with the highest evidence from Set 2. The most likely models are ABS from Set 1 and BB type from Set 2. The model evidence for each aerosol model from Set 1 is shown in Figure 10. Note that the differences between the model evidences for the four aerosol models are not very large, and the model evidence of ABS was slightly larger than those of the other models. In Set 2, there are 26 aerosol models and five aerosol types. The sum of the first 10 best aerosol model evidences for each aerosol type from Set 2 are illustrated in Figure 11. The most probable type is BB. In conclusion, the most plausible aerosol models identified by the algorithm, that is, ABS from Set 1 and BB from Set 2, are of the biomass burning aerosol type. This strongly absorbing aerosol type is consistent with the thick smoke observed in the true-color image.

The predominant models for retrieval on 5 July 2019 are ABS and DUST among Set 1 (Figure 12a), and BB and DD among Set 2 (Figure 12b). Thus, in addition to the aerosol models identified for 4 July 2019, the dust aerosol model comes into play. The model evidence for each aerosol model from Set 1 and each aerosol type from Set 2 are displayed in Figures 13 and 14, respectively. In conclusion, compared with that on 4 July 2019, the dominance of biomass burning aerosol type (ABS and BB) is less obvious. Taking into account the thinner smoke on 5 July 2019 and the long traveling distance from the origin, the presence of a less absorbing mixture of different aerosol types (biomass burning and dust) seems to be plausible. The corresponding maximum and mean solution estimates are shown in Figures 15–18 (Figures 15 and 17 for Set 1, and Figures 16 and 18 for Set 2).

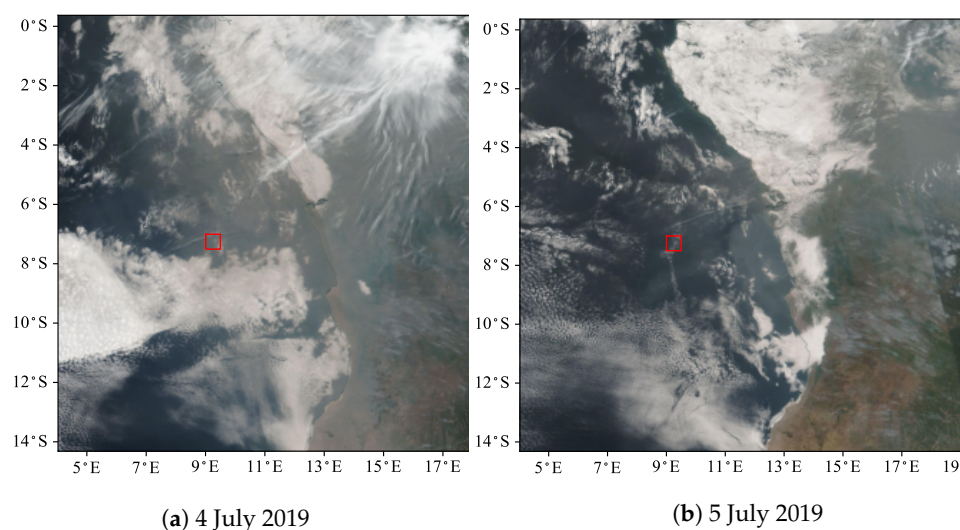


Figure 8. True-color VIIRS images recorded on (a) 4 and (b) 5 July 2019.

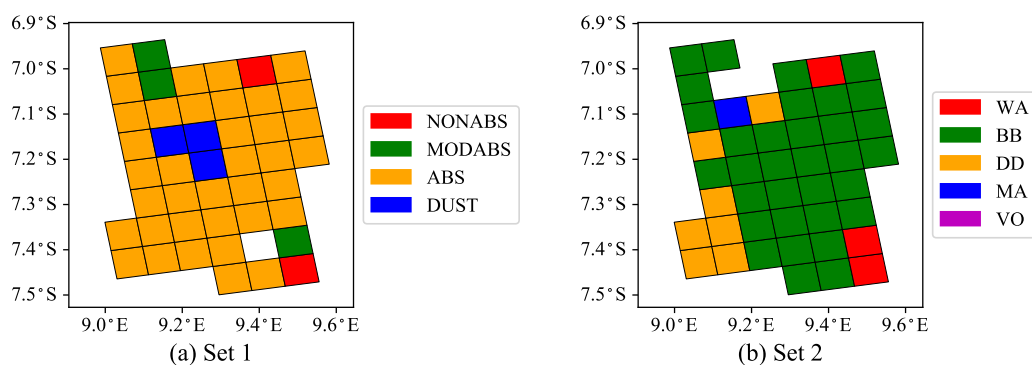


Figure 9. (a) The aerosol model with the highest evidence from Set 1, and (b) the aerosol type containing the aerosol model with the highest evidence from Set 2. The TROPOMI spectra were recorded on 4 July 2019.

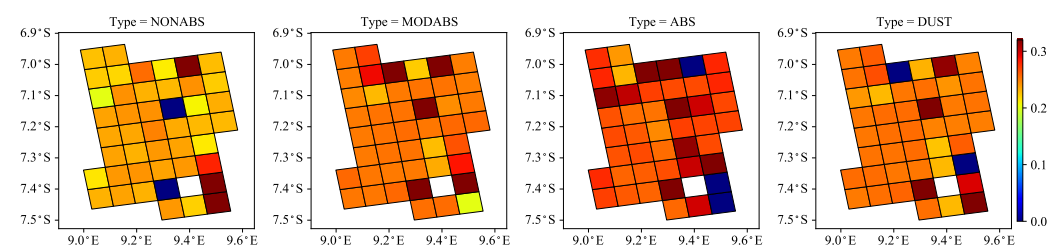


Figure 10. The model evidence for each aerosol model from Set 1. The TROPOMI spectra were recorded on 4 July 2019.

To demonstrate the performance under various aerosol loading scenarios, we performed retrievals for another two cases from TROPOMI. The first case focused on a desert dust aerosol case in Sahara on 6 June 2020 (see Figure S1 for the VIIRS image). The model evidence for the aerosol models in Set 1 and the aerosol types of Set 2 are given in Figures S2–S4. The prevailing aerosol model and aerosol type are DUST from Set 1 and DD from Set 2, given the fact that both models represent desert dust aerosols. The second case was for a urban aerosol case on 10 February 2020 over eastern China (see Figure S7 for the VIIRS image) where many industrial cities are located. As shown in Figures S8–S10, the NONABS model in Set 1 and the WA aerosol type in Set 2 are the most plausible choices, as both stands for industrial aerosols.

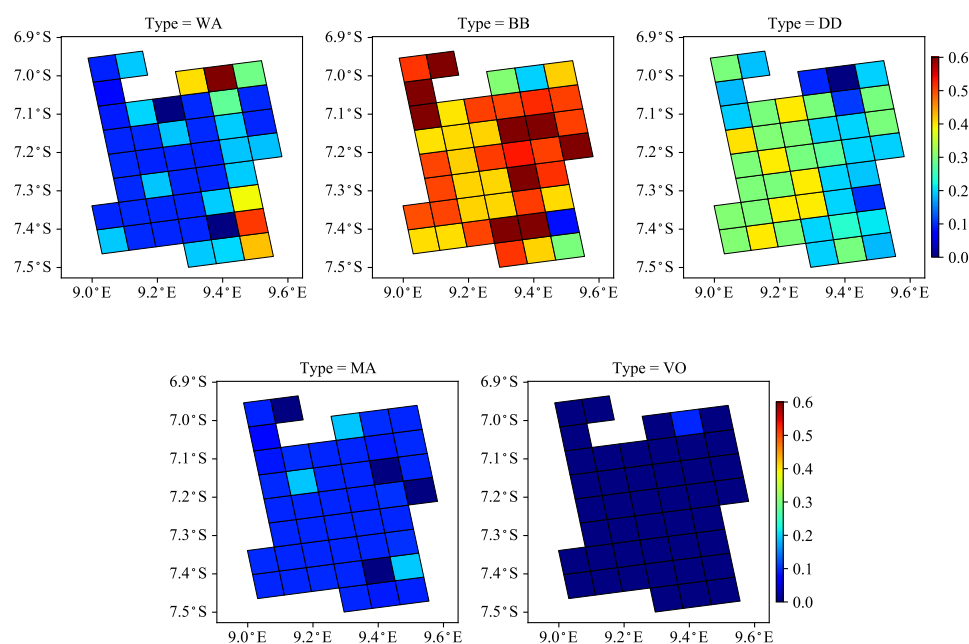


Figure 11. The sum of the first 10 best aerosol model evidences for each aerosol type from Set 2. The TROPOMI spectra were recorded on 4 July 2019.

Figures S5 and S6 illustrate the maximum and mean solution estimates for the first case, respectively. Figures S11 and S12 depict the corresponding solution estimates for the second case.

The dominant aerosol type or aerosol model for each study can be found from the above analysis. However, the most likely model varies from pixel to pixel, indicating that sometimes a “wrong” model may be chosen, which is consistent with the findings using the synthetic data. Based on the results of the retrieval solutions, we can see that

1. the mean solution estimates show a slightly smoother spatial pattern than the maximum solution estimates, and
2. despite the differences in the micro-physical properties of the aerosol models from Sets 1 and 2, the spatial distributions of the mean retrieval results are comparable.

In this study, the state vector was a two-element vector (aerosol optical depth and layer height) by considering the box profile for simplicity. The degree of freedom was estimated to be 2 in most cases. From the practical point of view, the retrieval on a Dell desktop (with 12 processors at 3.2 GHz, 31.2 GB of RAM) took less than 10 min (10–15 iterations in total) by running the program with all the models included in Set 1 and longer than 60 min (approximately 100 iterations in total) by considering all the models included in Set 2.

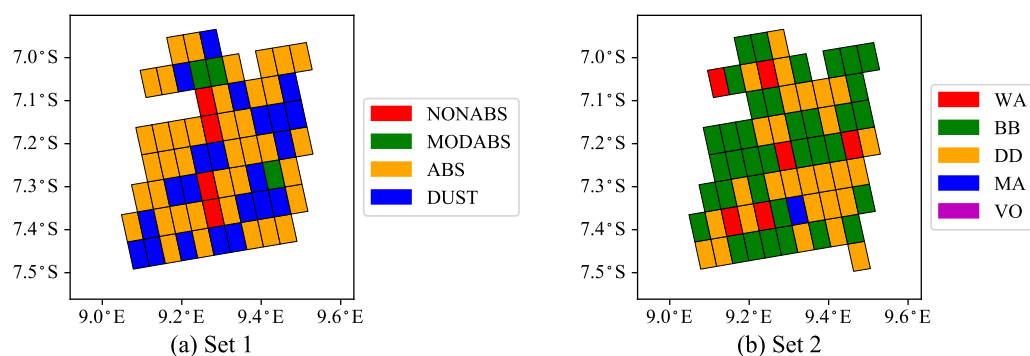


Figure 12. The same as in Figure 9 but for the data on 5 July 2019.

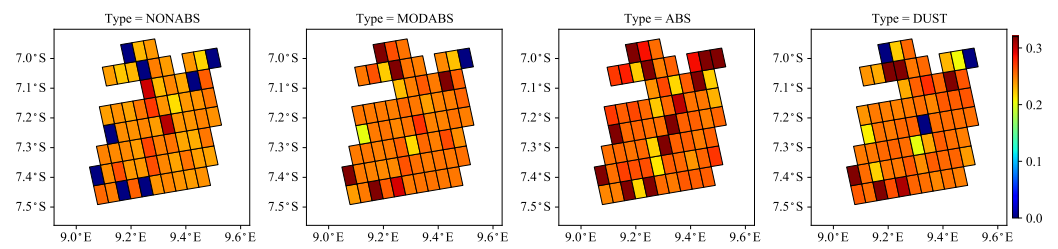


Figure 13. The same as in Figure 10 but for the data on 5 July 2019.

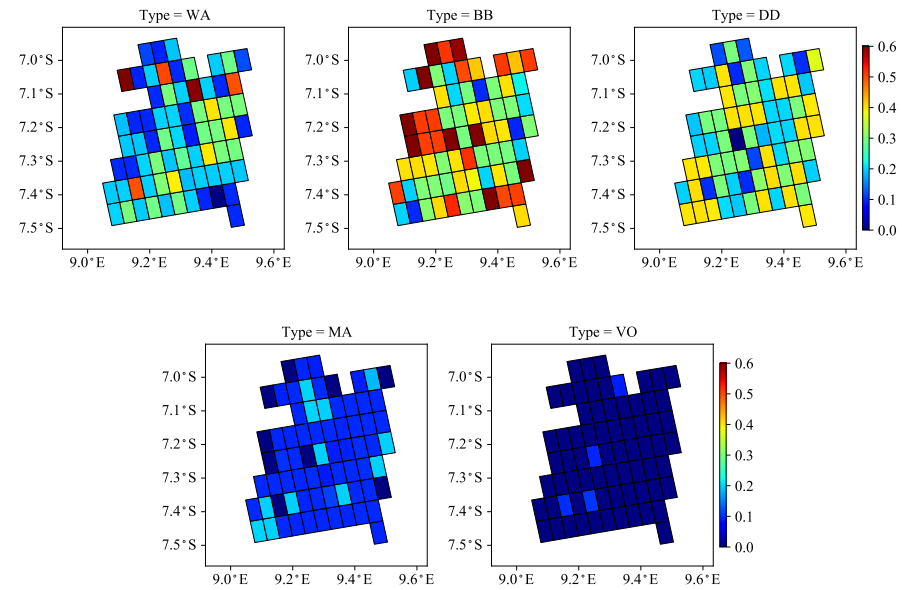


Figure 14. The same as in Figure 11 but for the data on 5 July 2019.

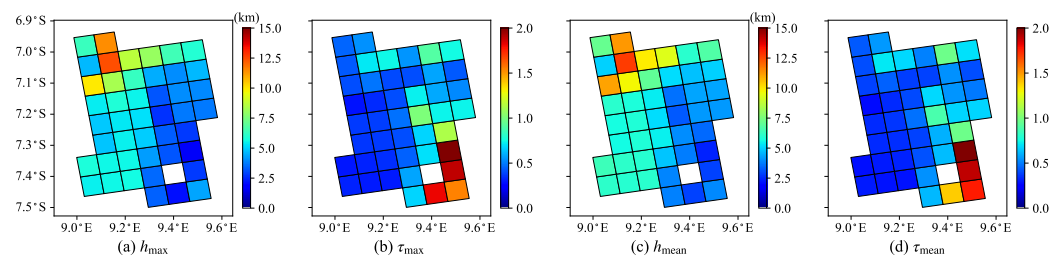


Figure 15. The maximum solution estimates (h_{\max} , τ_{\max}) and the mean solution estimates (h_{mean} , τ_{mean}) for Set 1 and data on 4 July 2019.

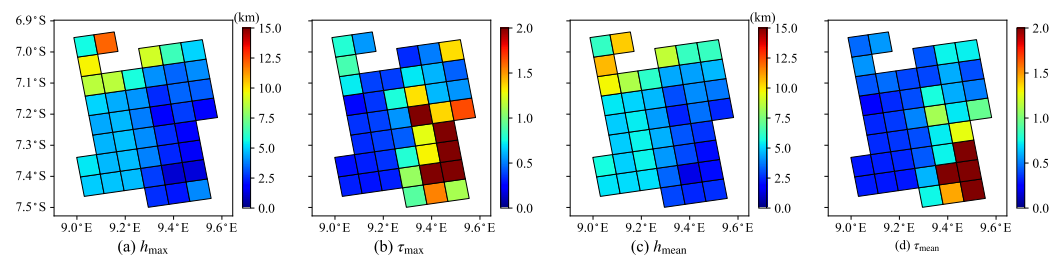


Figure 16. The maximum solution estimates (h_{\max} , τ_{\max}) and the mean solution estimates (h_{mean} , τ_{mean}) for Set 2 and data on 4 July 2019.

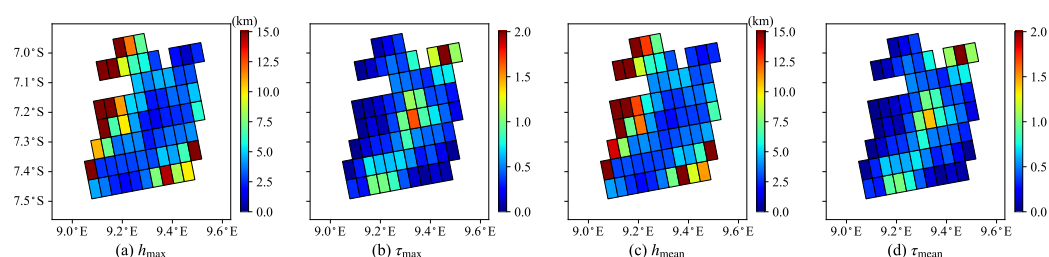


Figure 17. The same as in Figure 15 but for the data on 5 July 2019.

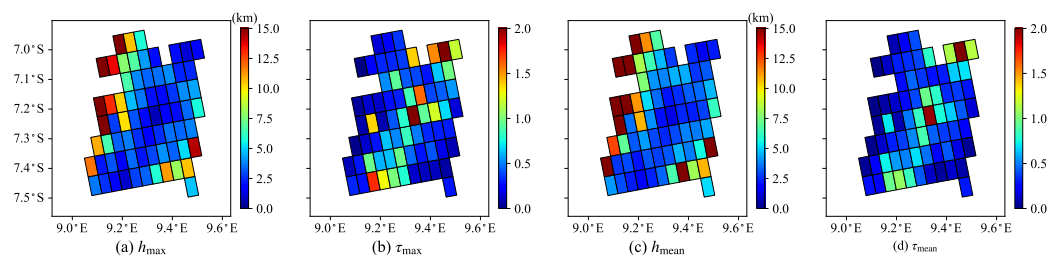


Figure 18. The same as in Figure 16 but for the data on 5 July 2019.

6. Conclusions

In this paper, the results of aerosol retrieval computed by means of a Bayesian algorithm that takes into account the uncertainty in aerosol model selection are presented. The solution corresponding to a specific aerosol model is characterized by a relative evidence which is used to construct (i) the maximum solution estimate, corresponding to the aerosol model with the highest evidence, and (ii) the mean solution estimate, representing a linear combination of solutions weighted by their evidences. The algorithm is applied to the retrieval of aerosol optical depth and layer height from synthetic and real TROPOMI data. The real TROPOMI data were taken on 4–5 July 2019 over a wild fire scene in South Africa. In the retrieval, two sets of aerosol models are taken into account; these correspond to the MODIS dark-target and OMAERO algorithms. The following conclusions are drawn.

1. When the exact aerosol model, for which synthetic data are generated, is included in the set of candidate models, the relative errors corresponding to the maximum solution estimate are relatively small. When this is not the case, it is likely that several aerosol models are able to fit the data equally well. In such situations, the mean solution estimate has a smaller bias than the maximum solution estimate.
2. For the real measurements on 4 July 2019, the absorbing aerosol model from Set 1 and the biomass burning aerosol type from Set 2 are found to be the most plausible. This result is in agreement with the thick smoke observed in the true-color image. For the thinner smoke scenario on 5 July 2019, the above models together with the dust aerosol model are found to be the most probable aerosol models. Actually, no dominant aerosol model, but rather a less absorbing mixture of different aerosol types, is identified in this case. The mean and maximum solution estimates have a similar spatial distribution, but the mean solution estimates have a more continuous spatial pattern.
3. The two TROPOMI cases on 6 June 2020 and 10 February 2020 for desert dust and urban aerosols, respectively, have demonstrated the promising performance of the proposed algorithm under various aerosol loading scenarios.
4. A definite choice between Sets 1 and 2 for possible candidate models may not exist and a suitable one could be problem dependent.

Note that when applying the Bayesian approach, we have to perform a retrieval for each candidate model. For this reason, the retrieval algorithm is computationally expensive, especially a set like Set 2 contains a larger number of aerosol models. To enhance its efficiency, development of a machine learning-based scheme is currently ongoing.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/rs13132489/s1>, Figure S1: True-color VIIRS image recorded on 6 June 2020. Figure S2: (a) The aerosol model with the highest evidence from Set 1, and (b) the aerosol type containing the aerosol model with the highest evidence from Set 2. The TROPOMI spectra were recorded on 6 June 2020. Figure S3: The model evidence for each aerosol model from Set 1. The TROPOMI spectra were recorded on 6 June 2020. Figure S4: The sum of the first 10 best aerosol model evidences for each aerosol type from Set 2. The TROPOMI spectra were recorded on 6 June 2020. Figure S5: The maximum solution estimates (h_{\max} , τ_{\max}) and the mean solution estimates (h_{mean} , τ_{mean}) for Set 1 and data on 6 June 2020. Figure S6: The same as in Figure S5 but for Set 2. Figure S7: True-color VIIRS image recorded on 10 February 2020. Figure S8: (a) The aerosol model with the highest evidence from Set 1, and (b) the aerosol type containing the aerosol model with the highest evidence from Set 2. The TROPOMI spectra were recorded on 10 February 2020. Figure S9: The model evidence for each aerosol model from Set 1. The TROPOMI spectra were recorded on 10 February 2020. Figure S10: The sum of the first 10 best aerosol model evidences for each aerosol type from Set 2. The TROPOMI spectra were recorded on 10 February 2020. Figure S11: The maximum solution estimates (h_{\max} , τ_{\max}) and the mean solution estimates (h_{mean} , τ_{mean}) for Set 1 and data on 10 February 2020. Figure S12: The same as in Figure S11 but for Set 2.

Author Contributions: Conceptualization, L.R. and J.X.; methodology, J.X. and A.D.; software, L.R.; investigation, D.S.E. and D.G.L.; data curation, L.R. and J.X.; writing—original draft preparation, L.R., J.X. and A.D.; writing—review and editing, D.S.E., D.G.L., and A.D.; visualization, L.R.; supervision, D.G.L. and A.D.; project administration, D.S.E. and D.G.L. All authors have read and agreed to the published version of the manuscript

Funding: This research was supported by the DLR programmatic [Nachwuchsgruppe “Retrieval der nächsten Generation”, 2,472,469]. The work of L. Rao was funded by the Chinese Scholarship Council.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors are grateful to DLR and KNMI for sharing the related information of TROPOMI/S5P.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

| | |
|---------|---|
| ALH | Aerosol Layer Height |
| AOD | Aerosol Optical Depth |
| GE_LER | Geometry-dependent Effective Lambertian Equivalent Reflectivity |
| MODIS | Moderate-resolution Imaging Spectroradiometer |
| NIR | Near-infrared |
| OMI | Ozone Monitoring Instrument |
| OMAERO | OMI Multi-wavelength |
| S5P | Sentinel-5 Precursor |
| TROPOMI | Tropospheric Monitoring Instrument |
| ABS | Absorbing (Set 1) |
| DUST | Desert dust (Set 1) |
| MODABS | Moderately absorbing (Set 1) |
| NONABS | Non-absorbing (Set 1) |
| BB | Biomass Burning (Set 2) |
| DD | Desert Dust (Set 2) |
| MA | Marine (Set 2) |
| VO | Volcanic (Set 2) |
| WA | Weakly absorbing (Set 2) |

References

- Holben, B.; Vermote, E.; Kaufman, Y.J.; Tanré, D.; Kalb, V. Aerosol retrieval over land from AVHRR data-application for atmospheric correction. *IEEE Trans. Geosci. Remote Sens.* **1992**, *30*, 212–222. [\[CrossRef\]](#)
- Levy, R.; Mattoo, S.; Munchak, L.; Remer, L.; Sayer, A.; Patadia, F.; Hsu, N. The Collection 6 MODIS aerosol products over land and ocean. *Atmos. Meas. Tech.* **2013**, *6*, 2989–3034. [\[CrossRef\]](#)
- Jackson, J.M.; Liu, H.; Laszlo, I.; Kondragunta, S.; Remer, L.A.; Huang, J.; Huang, H.C. Suomi-NPP VIIRS aerosol algorithms and data products. *J. Geophys. Res. Atmos.* **2013**, *118*, 12673–12689. [\[CrossRef\]](#)
- Wei, J.; Li, Z.; Sun, L.; Peng, Y.; Zhang, Z.; Li, Z.; Su, T.; Feng, L.; Cai, Z.; Wu, H. Evaluation and uncertainty estimate of next-generation geostationary meteorological Himawari-8/AHI aerosol products. *Sci. Total Environ.* **2019**, *692*, 879–891. [\[CrossRef\]](#) [\[PubMed\]](#)
- Kahn, R.A.; Li, W.H.; Moroney, C.; Diner, D.J.; Martonchik, J.V.; Fishbein, E. Aerosol source plume physical characteristics from space-based multiangle imaging. *J. Geophys. Res. Atmos.* **2007**, *112*. [\[CrossRef\]](#)
- Virtanen, T.; Kolmonen, P.; Rodríguez, E.; Sogacheva, L.; Sundström, A.M.; de Leeuw, G. Ash plume top height estimation using AATSR. *Atmos. Meas. Tech.* **2014**, *7*, 2437–2456. [\[CrossRef\]](#)
- Dubovik, O.; Herman, M.; Holdak, A.; Lapyonok, T.; Tanré, D.; Deuzé, J.; Ducos, F.; Sinyuk, A.; Lopatin, A. Statistically optimized inversion algorithm for enhanced retrieval of aerosol properties from spectral multi-angle polarimetric satellite observations. *Atmos. Meas. Tech.* **2011**, *4*, 975–1018. [\[CrossRef\]](#)
- Nanda, S.; Graaf, M.d.; Veefkind, J.P.; Sneep, M.; Linden, M.t.; Sun, J.; Levelt, P.F. A first comparison of TROPOMI aerosol layer height (ALH) to CALIOP data. *Atmos. Meas. Tech.* **2020**, *13*, 3043–3059. [\[CrossRef\]](#)
- Kokhanovsky, A.; Davis, A.; Cairns, B.; Dubovik, O.; Hasekamp, O.; Sano, I.; Mukai, S.; Rozanov, V.; Litvinov, P.; Lapyonok, T.; et al. Space-based remote sensing of atmospheric aerosols: The multi-angle spectro-polarimetric frontier. *Earth Sci. Rev.* **2015**, *145*, 85–116. [\[CrossRef\]](#)
- Hess, M.; Koepke, P.; Schult, I. Optical properties of aerosols and clouds: The software package OPAC. *Bull. Am. Meteorol. Soc.* **1998**, *79*, 831–844. [\[CrossRef\]](#)
- Levy, R.C.; Remer, L.A.; Dubovik, O. Global aerosol optical properties and application to Moderate Resolution Imaging Spectroradiometer aerosol retrieval over land. *J. Geophys. Res. Atmos.* **2007**, *112*. [\[CrossRef\]](#)
- Torres, O.; Tanskanen, A.; Veihelmann, B.; Ahn, C.; Braak, R.; Bhartia, P.K.; Veefkind, P.; Levelt, P. Aerosols and surface UV products from Ozone Monitoring Instrument observations: An overview. *J. Geophys. Res. Atmos.* **2007**, *112*. [\[CrossRef\]](#)
- Chin, M.; Ginoux, P.; Kinne, S.; Torres, O.; Holben, B.N.; Duncan, B.N.; Martin, R.V.; Logan, J.A.; Higurashi, A.; Nakajima, T. Tropospheric aerosol optical thickness from the GOCART model and comparisons with satellite and Sun photometer measurements. *J. Atmos. Sci.* **2002**, *59*, 461–483. [\[CrossRef\]](#)
- Taylor, M.; Kazadzis, S.; Amiridis, V.; Kahn, R. Global aerosol mixtures and their multiyear and seasonal characteristics. *Atmos. Environ.* **2015**, *116*, 112–129. [\[CrossRef\]](#)
- Martonchik, J.; Diner, D.; Kahn, R.; Ackerman, T.; Verstraete, M.; Pinty, B.; Gordon, H. Techniques for the retrieval of aerosol properties over land and ocean using multiangle imaging. *IEEE Trans. Geosci. Remote Sens.* **1998**, *36*, 1212–1227. [\[CrossRef\]](#)
- Bellouin, N.; Boucher, O.; Tanré, D.; Dubovik, O. Aerosol absorption over the clear-sky oceans deduced from POLDER-1 and AERONET observations. *Geophys. Res. Lett.* **2003**, *30*. [\[CrossRef\]](#)
- Kim, J.; Lee, J.; Lee, H.C.; Higurashi, A.; Takemura, T.; Song, C.H. Consistency of the aerosol type classification from satellite remote sensing during the Atmospheric Brown Cloud–East Asia Regional Experiment campaign. *J. Geophys. Res. Atmos.* **2007**, *112*. [\[CrossRef\]](#)
- Penning de Vries, M.J.M.; Beirle, S.; Hörmann, C.; Kaiser, J.W.; Stammes, P.; Tilstra, L.G.; Tuinder, O.N.E.; Wagner, T. A global aerosol classification algorithm incorporating multiple satellite data sets of aerosol and trace gas abundances. *Atmos. Chem. Phys.* **2015**, *15*, 10597–10618. [\[CrossRef\]](#)
- Kumar, K.R.; Kang, N.; Yin, Y. Classification of key aerosol types and their frequency distributions based on satellite remote sensing data at an industrially polluted city in the Yangtze River Delta, China. *Int. J. Climatol.* **2018**, *38*, 320–336. [\[CrossRef\]](#)
- Hoeting, J.A.; Madigan, D.; Raftery, A.E.; Volinsky, C.T. Bayesian model averaging: A tutorial. *Stat. Sci.* **1999**, *14*, 382–401.
- Määttä, A.; Laine, M.; Tamminen, J.; Veefkind, J. Quantification of uncertainty in aerosol optical thickness retrieval arising from aerosol microphysical model and other sources, applied to Ozone Monitoring Instrument (OMI) measurements. *Atmos. Meas. Tech.* **2014**, *7*, 1185–1199. [\[CrossRef\]](#)
- Kauppi, A.; Kolmonen, P.; Laine, M.; Tamminen, J. Aerosol-type retrieval and uncertainty quantification from OMI data. *Atmos. Meas. Tech.* **2017**, *10*. [\[CrossRef\]](#)
- Sasi, S.; Natraj, V.; Molina García, V.; Efremenko, D.S.; Loyola, D.; Doicu, A. Model Selection in Atmospheric Remote Sensing with an Application to Aerosol Retrieval from DSCOVR/EPIC, Part 1: Theory. *Remote Sens.* **2020**, *12*, 3724. [\[CrossRef\]](#)
- Sasi, S.; Natraj, V.; Molina García, V.; Efremenko, D.S.; Loyola, D.; Doicu, A. Model Selection in Atmospheric Remote Sensing with Application to Aerosol Retrieval from DSCOVR/EPIC. Part 2: Numerical Analysis. *Remote Sens.* **2020**, *12*, 3656. [\[CrossRef\]](#)
- Marshak, A.; Herman, J.; Adam, S.; Karin, B.; Carn, S.; Cede, A.; Geogdzhayev, I.; Huang, D.; Huang, L.K.; Knyazikhin, Y.; et al. Earth observations from DSCOVR EPIC instrument. *Bull. Am. Meteorol. Soc.* **2018**, *99*, 1829–1850. [\[CrossRef\]](#)

26. Veefkind, J.; Aben, I.; McMullan, K.; Förster, H.; de Vries, J.; Otter, G.; Claas, J.; Eskes, H.; de Haan, J.; Kleipool, Q.; et al. TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications. *Remote Sens. Environ.* **2012**, *120*, 70–83. [\[CrossRef\]](#)
27. Spurr, R. LIDORT and VLIDORT. Linearized pseudo-spherical scalar and vector discrete ordinate radiative transfer models for use in remote sensing retrieval problems. In *Light Scattering Reviews*; Kokhanovsky, A., Ed.; Springer: Berlin/Heidelberg, Germany, 2008; Volume 3, pp. 229–271.
28. Efremenko, D.; Doicu, A.; Loyola, D.; Trautmann, T. Acceleration techniques for the discrete ordinate method. *J. Quant. Spectrosc. Radiat. Transf.* **2013**, *114*, 73–81. [\[CrossRef\]](#)
29. Goody, R.; West, R.; Chen, L.; Crisp, D. The correlated-k method for radiation calculations in nonhomogeneous atmospheres. *J. Quant. Spectrosc. Radiat. Transf.* **1989**, *42*, 437–651. [\[CrossRef\]](#)
30. Natraj, V.; Shia, R.L.; Huang, X.; Margolis, J.S.; Yung, Y.L. Application of principal component analysis to high spectral resolution radiative transfer: A case study of the O₂ A band. *J. Quant. Spectrosc. Radiat. Transf.* **2005**, *95*, 539–556. [\[CrossRef\]](#)
31. Natraj, V.; Shia, R.L.; Yung, Y.L. On the use of principal component analysis to speed up radiative transfer calculations. *J. Quant. Spectrosc. Radiat. Transf.* **2010**, *111*, 810–816. [\[CrossRef\]](#)
32. Xu, J.; Schreier, F.; Doicu, A.; Trautmann, T. Assessment of Tikhonov-type regularization methods for solving atmospheric inverse problems. *J. Quant. Spectrosc. Radiat. Transf.* **2016**, *184*, 274–286. [\[CrossRef\]](#)
33. Xu, J.; Rao, L.; Schreier, F.; Efremenko, D.S.; Doicu, A.; Trautmann, T. Insight into Construction of Tikhonov-Type Regularization for Atmospheric Retrievals. *Atmosphere* **2020**, *11*, 1052. [\[CrossRef\]](#)
34. Rao, L.; Xu, J.; Efremenko, D.S.; Loyola, D.G.; Doicu, A. Hyperspectral Satellite Remote Sensing of Aerosol Parameters. In Preparation.
35. Molina García, V.; Sasi, S.; Efremenko, D.S.; Doicu, A.; Loyola, D. Radiative transfer models for retrieval of cloud parameters from EPIC/DSCOVR measurements. *J. Quant. Spectrosc. Radiat. Transf.* **2018**, *213*, 228–240. [\[CrossRef\]](#)
36. Molina García, V.; Sasi, S.; Efremenko, D.S.; Doicu, A.; Loyola, D. Linearized radiative transfer models for retrieval of cloud parameters from EPIC/DSCOVR measurements. *J. Quant. Spectrosc. Radiat. Transf.* **2018**, *213*, 241–251. [\[CrossRef\]](#)
37. Patterson, H.D.; Thompson, R. Recovery of inter-block information when block sizes are unequal. *Biometrika* **1971**, *58*, 545–554. [\[CrossRef\]](#)
38. Smyth, G.K.; Verbyla, A.P. A conditional likelihood approach to residual maximum likelihood estimation in generalized linear models. *J. R. Stat. Soc. Ser. B Methodol.* **1996**, *58*, 565–572. [\[CrossRef\]](#)
39. Stuart, A.; Ord, K.; Arnold, S. *Kendall's Advanced Theory of Statistics. Volume 2A: Classical Inference and the Linear Model*; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2000.
40. Wahba, G. Practical approximate solutions to linear operator equations when the data are noisy. *SIAM J. Numer. Anal.* **1977**, *14*, 651–667. [\[CrossRef\]](#)
41. Wahba, G. *Spline Models for Observational Data*; SIAM: Philadelphia, PA, USA, 1990.
42. Meng, Z.; Yang, P.; Kattawar, G.W.; Bi, L.; Liou, K.; Laszlo, I. Single-scattering properties of tri-axial ellipsoidal mineral dust aerosols: A database for application to radiative transfer calculations. *J. Aerosol Sci.* **2010**, *41*, 501–512. [\[CrossRef\]](#)
43. Alados-Arboledas, L.; Müller, D.; Guerrero-Rascado, J.L.; Navas-Guzmán, F.; Pérez-Ramírez, D.; Olmo, F.J. Optical and microphysical properties of fresh biomass burning aerosol retrieved by Raman lidar, and star-and sun-photometry. *Geophys. Res. Lett.* **2011**, *38*. [\[CrossRef\]](#)
44. Waquet, F.; Cornet, C.; Deuzé, J.L.; Dubovik, O.; Ducos, F.; Goloub, P.; Herman, M.; Lapyonok, T.; Labonnote, L.C.; Riedi, J.; et al. Retrieval of aerosol microphysical and optical properties above liquid clouds from POLDER/PARASOL polarization measurements. *Atmos. Meas. Tech.* **2013**, *6*, 991–1016. [\[CrossRef\]](#)
45. Yu, X.; Shi, C.; Ma, J.; Zhu, B.; Li, M.; Wang, J.; Yang, S.; Kang, N. Aerosol optical properties during firework, biomass burning and dust episodes in Beijing. *Atmos. Environ.* **2013**, *81*, 475–484. [\[CrossRef\]](#)
46. Loyola, D.G.; Xu, J.; Heue, K.P.; Zimmer, W. Applying FP_ILM to the retrieval of geometry-dependent effective Lambertian equivalent reflectivity (GE_LER) daily maps from UVN satellite measurements. *Atmos. Meas. Tech.* **2020**, *13*, 985–999. [\[CrossRef\]](#)